Research Article

Journal of Molecular Microbiology and Biotechnology

J Mol Microbiol Biotechnol 2014;24:19–36
DOI: 10.1159/000354322

Published online: October 18, 2013

Analysis of the Complete Genomes of *Acholeplasma brassicae*, *A. palmae* and *A. laidlawii* and Their Comparison to the Obligate Parasites from '*Candidatus* Phytoplasma'

Michael Kube[a] Christin Siewert[a] Alexander M. Migdoll[c] Bojan Duduk[g] Sabine Holz[a] Ralf Rabus[d] Erich Seemüller[e] Jelena Mitrovic[g] Ines Müller[b, f] Carmen Büttner[a] Richard Reinhardt[b, f]

[a]Division Phytomedicine, Department of Crop and Animal Sciences, Humboldt-Universität zu Berlin, and [b]Max Planck Institute for Molecular Genetics, Berlin, [c]National Center for Tumor Diseases (NCT) Heidelberg, Heidelberg, [d]Institute for Chemistry and Biology of the Marine Environment, Carl von Ossietzky University of Oldenburg, Oldenburg, [e]Julius Kuehn Institute, Federal Research Centre for Cultivated Plants, Institute for Plant Protection in Fruit Crops and Viticulture, Dossenheim, and [f]Max Planck Genome Centre Cologne, Cologne, Germany; [g]Institute of Pesticides and Environmental Protection, Belgrade, Serbia

Key Words

Complete genomes · *Acholeplasma palmae* · *Acholeplasma brassicae* · *Candidatus* phytoplasma

Abstract

Analysis of the completely determined genomes of the plant-derived *Acholeplasma brassicae* strain O502 and *A. palmae* strain J233 revealed that the circular chromosomes are 1,877,792 and 1,554,229 bp in size, have a G + C content of 36 and 29%, and encode 1,690 and 1,439 proteins, respectively. Comparative analysis of these sequences and previously published genomes of *A. laidlawii* strain PG-8, '*Candidatus* Phytoplasma asteris' strains, '*Ca*. P. australiense' and '*Ca*. P. mali' show a limited shared basic genetic repertoire. The acholeplasma genomes are characterized by a low number of rearrangements, duplication and integration events. Exceptions are the unusual duplication of rRNA operons in *A. brassicae* and an independently introduced second gene for a single-stranded binding protein in both genera. In contrast to phytoplasmas, the acholeplasma genomes differ by encoding the cell division protein FtsZ, a wide variety of ABC transporters, the F_0F_1 ATP synthase, the *Rnf*-complex, SecG of the *Sec*-dependent secretion system, a richly equipped repertoire for carbohydrate metabolism, fatty acid, isoprenoid and partial amino acid metabolism. Conserved metabolic proteins encoded in phytoplasma genomes such as the malate dehydrogenase SfcA, several transporters and proteins involved in host-interaction, and virulence-associated effectors were not predicted for the acholeplasmas.

Introduction

Acholeplasmas are cell wall-less bacteria belonging to the class Mollicutes. The members of the order Acholeplasmatales do not require sterol for growth as indicated by the name [Saito et al., 1977] and were therefore separated from the related order Mycoplasmatales. Achole-

C.S. and A.M.M. contributed equally to this work.

KARGER

© 2013 S. Karger AG, Basel
1464–1801/14/0241–0019$39.50/0

E-Mail karger@karger.com
www.karger.com/mmb

Michael Kube
Humboldt-Universität zu Berlin, Department of Crop and Animal Sciences
Division Phytomedicine
Lentzeallee 55/57, DE–14195 Berlin (Germany)
E-Mail Michael.Kube@agrar.hu-berlin.de

plasmas colonize a wide variety of habitats as saprophytes and have been described as commensals in vertebrates, insects and plants. Two *Acholeplasma* species *(A. brassicae* and *A. palmae)* were named according to their isolation from broccoli *(Brassica oleracea* var. *italica)* and a coconut tree *(Cocos nucifera),* respectively [Tully et al., 1994]. Other *Acholeplasma* species such as *A. laidlawii*, *A. axanthum* and *A. oculi* were also detected on plant surfaces [Brown et al., 2011]. In addition, acholeplasmas including *A. pleciae, A. laidlawii* and *A. morum* were identified in pools of insects such as *Anopheles sinensis* and *Armigeres subalbatus* and may multiply in phytoplasma vectors [Edengreen and Markham, 1987]. However, no acholeplasma primary pathogen has been described to date. 16S-rDNA-based phylogenetic analysis indicates that the genus *Acholeplasma* is the most closely related taxon of the provisional monophyletic genus '*Candidatus* Phytoplasma' [IRPCM, 2004; Lee et al., 2000]. It has been suggested that phytoplasmas and acholeplasmas come from a (*Acholeplasma*-like last) common ancestor [Zhao et al., 2009]. In this lineage, *Acholeplasma* species such as *A. laidlawii* show a deeper branching than phytoplasmas [Ogawa et al., 2011]. The phylogenetic differences are also reflected by the fact that phytoplasmas separate from the saprophytic acholeplasmas by their association to many plant diseases including that of important crops [Strauss, 2009]. Phloem-sucking insect vectors mainly spread phytoplasmas. They represent obligate parasites, restricted in plants to the phloem sieve tubes. The spread of these pests is supported by a manipulation of the plant hosts and insect vectors by the secretion of phytoplasma proteins [Sugio et al., 2011].

Both genera have small genomes of about 1.2–2.1 Mb in size with a G + C content of 27–38% for acholeplasmas [Carle et al., 1995; Neimark and Kirkpatrick, 1993] and 0.5–1.4 Mb with a G + C content of 21–33% for phytoplasmas [IRPCM, 2004; Kube et al., 2008; Marcone et al., 1999; Marcone and Seemüller, 2001; Neimark and Kirkpatrick, 1993]. They share features such as the usage of the bacterial genetic code, including the regular usage of UGA to encode a termination signal, and differ in this respect from the genera *Mycoplasma*, *Ureaplasma* and *Spiroplasma* [Razin et al., 1998]. Genomes of four phytoplasmas were completely determined. They comprise '*Ca*. P. asteris' strains OY-M and AY-WB [Bai et al., 2006; Oshima et al., 2004], '*Ca*. P. australiense' [Tran-Nguyen et al., 2008] and '*Ca*. P. mali' [Kube et al., 2008]. For the acholeplasmas, only the complete genome sequence of *A. laidlawii* strain PG-8A was determined [Lazarev et al., 2011]. This genome consists of a single circular chromosome with a size of 1,497 kb and a G + C content of 31%, and represented the largest chromosome among the completely sequenced Mollicutes. Contrasting other Mollicutes, Lazarev et al. [2011] revealed for *A. laidlawii* a richly equipped repertoire for metabolism, SOS response, repair systems and extensive genetic equipment for transcriptional regulation including the two-component systems, riboswitches and T-boxes. The authors interpreted this genetic repertoire as necessary for the adaptation to changing environmental conditions. Environmental stresses are more relevant for acholeplasmas than for phytoplasmas due to their parasitic lifestyle and host dependency. In contrast, the metabolism of *Acholeplasma* and other Mollicutes depends on external sources [Razin, 1978]. Carbohydrate metabolism, and in particular glycolysis, was assigned as the only pathway to generate ATP in *A. laidlawii* as in other fermenting Mollicutes [Lazarev et al., 2011]. A few amino acids such as phenylalanine and tyrosine can be generated de novo, but the majority has to be imported. A similar situation applies for the cofactors, vitamins and nucleotide metabolism that are partially encoded and offer the genetic repertoire for conversion of intermediates. Pathways for the de novo biosynthesis of carotenoids, fatty acids and lipids are encoded in *A. laidlawii*. These features also separate *Acholeplasma* species from the phytoplasmas. A proteomic survey resulted in the identification of 58% of the predicted proteins of *A. laidlawii* and provided insights into the post-translational modification by phosphorylation and acylation of proteins, confirming the prevalence on palmitic acid acylation [Lazarev et al., 2011], which is also known from various *Mycoplasma* species [Worliczek et al., 2007]. Lazarev et al. [2011] concluded that the encoded capabilities and known ability of *A. laidlawii* to adapt to various environments indicate that the acholeplasmas form a unique branch of evolution. The authors therefore suggested that acholeplasmas could not be interpreted as an intermediate in genome condensation or side trend in the evolution of parasitism. This assumption considers the parasitic lifestyle and loss of major metabolic pathways of the phytoplasmas [Kube et al., 2012].

Many questions concerning the evolution of acholeplasmas and phytoplasmas remain open. The present study addresses some of them. The *Acholeplasma* species differ by up to one third in their genome size. It is unclear whether this is due to genome condensation and adaptation to the environment or due to duplication of genetic material and integration events as it is known for phytoplasmas, and whether *A. laidlawii* is representative of its genus. The close phylogenetic relationship to the phyto-

Table 1. Genome features of complete acholeplasma genomes in comparison to phytoplasmas

	Acholeplasma			*Ca.* Phytoplasma			
	brassicae O502	*palmae* J233	*laidlawii* PG-8A	asteris OY-M	asteris AY-WB	australiense Rp-A	mali AT
Chromosome organization	circular	circular	circular	circular	circular	circular	linear
Chromosome size, bp	1,877,792	1,554,229	1,496,992	853,092	706,569	879,959	601,943
G + C content, %	35.77	28.98	31.93	27.76	26.89	27.42	21.39
G + C % of protein coding genes[1]	36.15	29.20	32.23	29.09	28.54	28.72	22.58
Protein-coding genes[1], n	1,690 (14)	1,439 (2)	1,380	752	671	684 (155)	481 (16)
Protein coding, %[1]	90.3	90.6	90.7	73.1	73.7	64.1	76.3
Average ORF size[1], nt	1,003	979	984	829	776	825	955
Protein coding genes/kb[1]	0.899	0.925	0.921	0.881	0.949	0.777	0.799
rRNA operons (n genes)	4 (14)[2]	2 (6)	2 (6)	2 (6)	2 (6)	2 (6)	2 (6)
G + C % of rRNA genes	49.31	48.64	48.57	45.95	46.14	46.37	44.32
tRNAs, n	45	35	34	32	31	35	32
G + C % of tRNA genes	56.12	56.56	56.97	53.61	53.64	54.11	52.41
Plasmids, n	–	–	–	2	4	1	–
Data source	FO681348 this study	FO681347 this study	CP000896.1 Lazarev et al., 2011	AP006628.2 Oshima et al., 2004	CP000061.1 Bai et al., 2006	AM422018.1 Tran-Nguyen et al., 2008	CU469464.1[3] Kube et al., 2008

[1] Genes assigned as pseudogenes are not incorporated. [2] *A. brassicae* encodes unusual rRNA operon structures including 5S-rRNA duplications. [3] CU469464 was recently updated.

plasmas also raises the question of whether plant-derived isolates of *A. brassicae* and *A. palmae* can provide information on an evolutionary adaptation to plant parasitism. For this reason, we determined and analyzed the complete genome sequences of *A. brassicae* strain O502 and *A. palmae* strain J233.

Results and Discussion

Genome Features

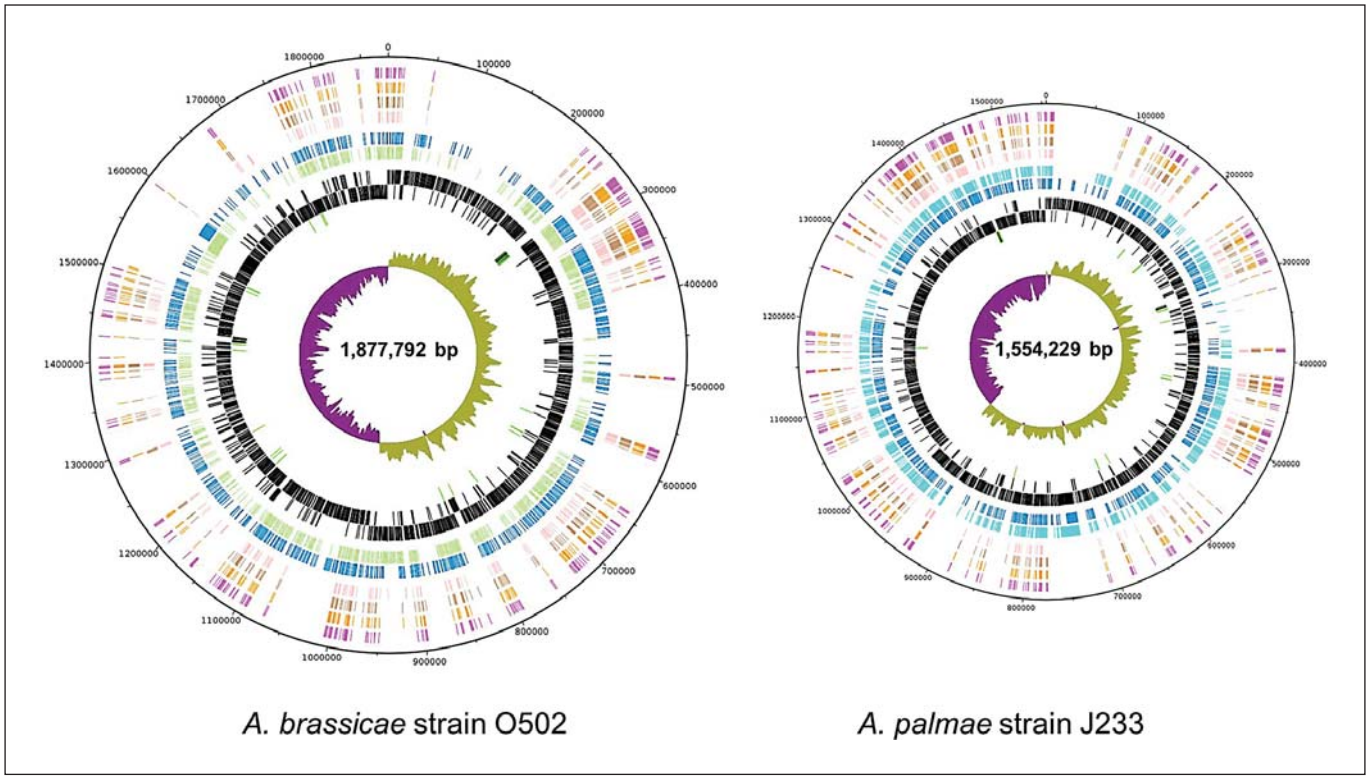
Complete genome sequences were obtained from 16,720 Sanger-derived sequences and 164,899 pyrosequencing-derived reads for *A. brassicae* strain O502 and 11,468 and 167,068, respectively, for *A. palmae* strain J233. A 26-fold sequencing coverage was obtained for *A. brassicae* and a 28-fold sequencing coverage was obtained for *A. palmae*.

The genetic content of the three acholeplasmas genomes is organized in single circular chromosomes. With a size of 1.9 Mb, *A. brassicae* possesses the largest known acholeplasma chromosome (table 1, fig. 1), while with 1.6 Mb *A. palmae* shows a genome size similar to that of *A. laidlawii* (1.5 Mb). The three genomes significantly differ in the G + C content of the chromosomes. The highest G + C content was observed for *A. brassicae* with 36%, whereas *A. palmae* has with 29% an even lower value than *A. laidlawii* (32%). The G + C content of *A. palmae* is similar to that of the '*Ca.* P. asteris' genomes (27–28%); with respect to the protein-encoding genes it is 29% for the three genomes.

The number of predicted proteins corresponds to the chromosome length, resulting in 1,690 for *A. brassicae*, 1,439 for *A. palmae* and 1,380 for *A. laidlawii*. The highest average *orf*-length of 1,003 nt was observed for *A. brassicae*, while the average value is 979 nt for *A. palmae* and 984 nt for *A. laidlawii*. These values are above the average *orf*-lengths (776–955 nt) of the complete phytoplasma genomes.

Numbers of shared proteins were estimated by reciprocal BLASTP (table 2). The highest number of shared orthologous proteins from deduced protein sets are found in *A. brassicae* and *A. laidlawii* (830 proteins), followed by *A. laidlawii* and *A. palmae* (765 proteins), corresponding to their phylogenetic positions. In contrast to *A. brassicae*, *A. palmae* and *A. laidlawii* belong to one major phylogenetic branch of the acholeplasmas [Volokhov et al., 2007].

In accordance with previous analysis of the genome content shared with *A. laidlawii* [Lazarev et al., 2011], relatively low numbers are obtained from the comparison of acholeplasmas and phytoplasmas (259–285 proteins), in-



Color version available online

Fig. 1. Comparative analyses of the deduced protein content of *A. brassicae*, *A. palmae*, *A. laidlawii*, '*Ca.* P. asteris' strain OY-M and AY-WB, '*Ca.* P. australiense', and '*Ca.* P. mali' via reciprocal BLASTP using *A. brassicae* or *A. palmae* as the reference. Circular patterns (from outside to inside): 1 (black circle), scale in base pairs of the reference chromosome; 2 (magenta), proteins of '*Ca.* P. australiense'; 3 (orange), proteins of '*Ca.* P. asteris' strain OY-M; 4 (brown), proteins of '*Ca.* P. asteris' strain AY-WB; 5 (light red), proteins of '*Ca.* P. mali'; 6/7 (dark blue), proteins of *A. laidlawii*; 6/7 (light green/light blue), proteins of *A. palmae* and *A. brassicae,* respectively; 8 and 9 (black), proteins of the reference genomes of *A. brassicae* or *A. palmae* by forward and reverse strand orientation; 9 (black and green), rRNA and tRNA of the reference genome and 9 (olive and pink), G+C skew. Colors refer to the online version only.

dicating the phylogenetic distance of both genera and the shared minimal core content, which is not scattered over the whole phytoplasma genomes (fig. 1).

Paralogs have little impact on gene content and chromosome length of the acholeplasma genomes in contrast to phytoplasmas (table 3). The *A. brassicae* chromosome encodes the highest percentage of paralogs (2.5%, 43 proteins) of the three acholeplasmas. Thus, gene duplication events have little impact on genome size. Its effect is much lower than in phytoplasmas where chromosome size increases in parallel with the number of repeated genes. '*Ca.* P. mali' represents a special case due to the terminal inverted repeat structures of the linear chromosome [Kube et al., 2008].

In *A. palmae* and *A. brassicae,* 16 and 28 genes are assigned to the prophages (online supplementary material 1; for all online supplementary material, see www.karger.com/doi/10.1159/000354322). One integrase (Abra_14190) shows 30% identity to the integrase of *Acholeplasma* phage L2 (P42540), which is organized in a circular double-stranded construct and described to integrate into the host genome [Maniloff et al., 1994]. In addition, the two genomes encode 24 and 48 putative transposases (including putative truncated genes), respectively, while only six transposases and two integrases are annotated in *A. laidlawii* [Lazarev et al., 2011]. Transposons associated with the so-called potential mobile units (PMUs) and phages are suggested to be the main driving forces in phytoplasmas (reviewed in [Kube et al., 2012]). PMU-like elements containing a similar genetic repertoire encoded by PMU1 [Toruno et al., 2010] could not be identified in the three acholeplasmas, but the PMU-associated Tra5-like transposases are encoded in both genera.

Table 2. Estimation of potential orthologous proteins by pairwise reciprocal BLASTP of chromosomal proteins (values in bold indicate high numbers obtained in acholeplasmas or between *A. palmae* and phytoplasmas)

	A. brassicae	*A. laidlawii*	*A. palmae*	'*Ca.* P. asteris' strain AY-WB	'*Ca.* P. asteris' strain OY-M	'*Ca.* P. australiense'	'*Ca.* P. mali'
A. brassicae	–	**830**	**742**	263	264	264	259
A. laidlawii		–	**765**	270	273	271	265
A. palmae			–	**283**	**285**	**285**	**282**
'*Ca.* P. asteris' strain AY-WB				–	402	355	296
'*Ca.* P. asteris' strain OY-M					–	344	295
'*Ca.* P. australiense'						–	296
'*Ca.* P. mali'							–

Several mobile genetic elements in bacteria are also characterized by a deviating codon usage indicating a horizontal gene transfers (HTG) [Waack et al., 2006], but also a so far not completed amelioration of the codon usage [Medrano-Soto et al., 2004]. The highest gene numbers, percentage and total length of this gene group is assigned to *A. palmae* (54, 3.8%, 35.7 kb), followed by *A. laidlawii* (29, 2.1%, 16.6 kb) and *A. brassicae* (17, 1%, 8.3 kb). Higher percentages of the total gene set are assigned to the phytoplasma chromosomes (3.9–5.1%), except for '*Ca.* P. mali' with only 0.6% (online suppl. material 2, 3). The majority of these genes has no assigned function except for several prophage elements and transposases (online suppl. material 1, 4). In summary, gene duplication and/or integration events have a minor impact on genome size in the three acholeplasmas in contrast to the phytoplasmas. However, limited gene duplications are present.

Table 3. Estimated paralogs in acholeplasma and phytoplasma chromosomes

Species	Protein-coding sequences[1]	Paralog genes		Σ paralog gene sequences (% of chromosome)	
		n	%	bp	%
A. brassicae	1,704	43	2.5	28,134	1.5
A. palmae	1,441	16	1.1	17,967	1.2
A. laidlawii	1,380	0	0	0	0
'*Ca.* P. australiense'	839	259	30.9	176,244	20.0
'*Ca.* P. asteris'					
Strain OY-M	752	155	20.6	117,339	13.8
Strain AY-WB	671	64	9.5	36,579	5.2
'*Ca.* P. mali'	497	69	13.9	59,838	9.9

[1] Annotated pseudogenes included.

rRNA Operon Organization

The genomes of *A. palmae*, *A. laidlawii* and the four completely determined phytoplasmas contain two rRNA operons, which are organized in the regular 16S-23S-5S-rRNA arrangement. In these genomes, one rRNA operon localizes downstream of the 5S-rRNA a tRNA cluster ($tRNA^{Val, Thr, Lys, Leu, Ala, Met, Met, Ser, Met, Asp, Phe}$) in a conserved order. *A. brassicae* deviates from Mollicutes by possessing two ribosomal units consisting of two rRNA operons each (fig. 2). They are organized as 16S-23S-5S-5S followed by a 16S-23S-5S-rRNA unit. No indication for a horizontal gene transfer was obtained. Unusual rRNA operon organizations are known from *M. fermentans* strains JER (CP001995) and M64 [Shu et al., 2011], *M. synoviae* strain 53 [Vasconcelos et al., 2005], *M. bovis* strain Hubei-1 (CP002513) and *M. pulmonis* strain UAB CTIP (AL445566) (online suppl. material 5). The last three strains mentioned possess one additional 5S-rRNA gene compared to the number of 16S-23S rRNA units.

The four complete phytoplasma genomes encode $tRNA^{Ile}$ in the 16S-23S-rRNA spacer region. *A. palmae* and *A. laidlawii* encode $tRNA^{Ile}$ and $tRNA^{Ala}$ in the rRNA operon on the reverse strand in contrast. All three acholeplasma genomes share one rRNA operon-organized 16S-23S-5S-rRNA and a downstream localized gene for $tRNA^{Asn}$. No tRNA encoding gene was detected between 16S- and 23S-rRNA on the forward strand for these two chromosomes. In contrast, *A. brassicae* encodes a $tRNA^{Ile}$ between the 16S- and 23S-rRNA genes of both operons of the first rRNA locus. It is remarkable for *A. brassicae* that the second rRNA operon of the first locus shows rRNA genes and tRNAs in the same conserved synteny also observed in phytoplasmas.

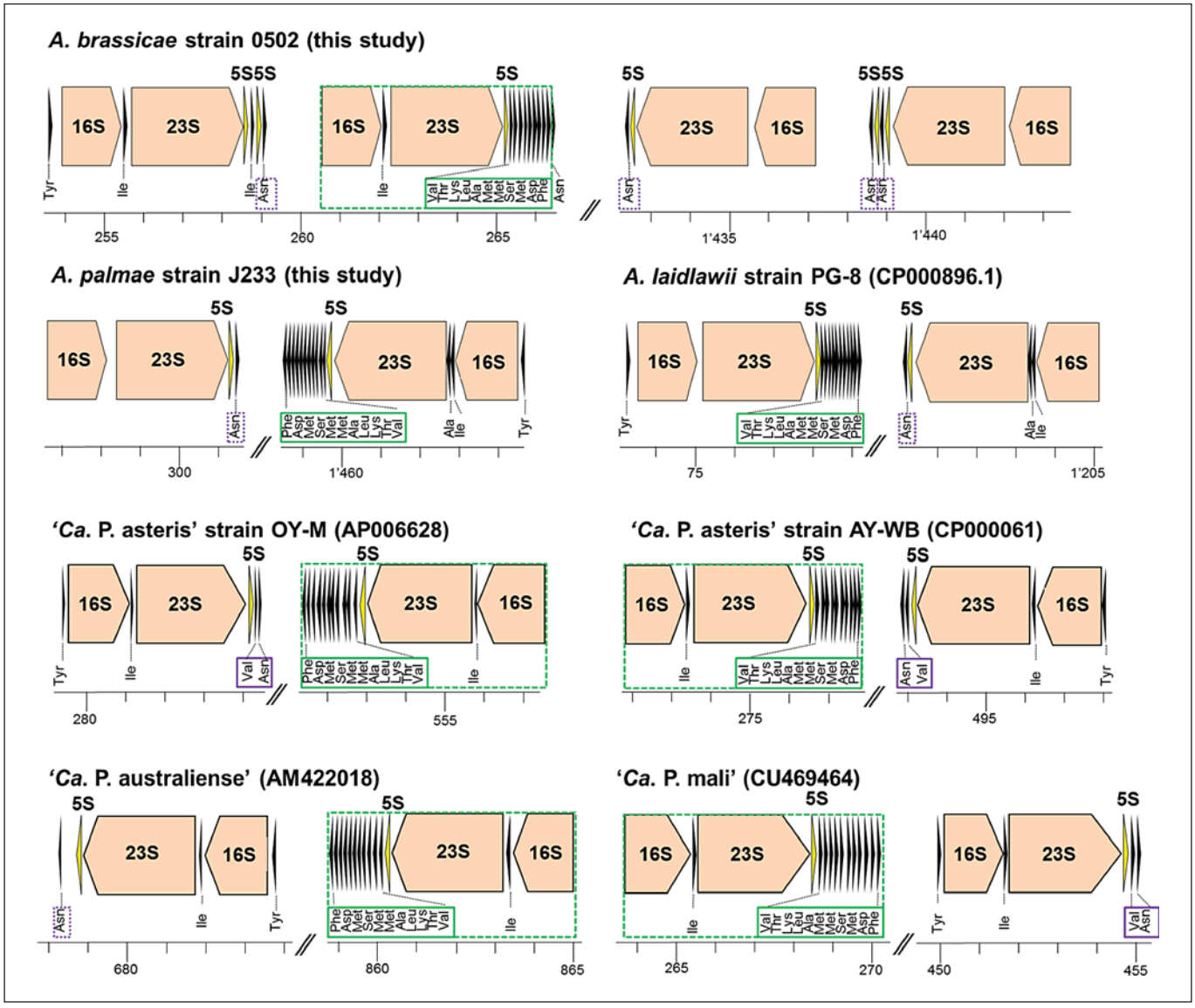


Fig. 2. rRNA operons and flanking tRNAs. Organization is shown for *A. brassicae* strain 0502, *A. palmae* strain J233 and *A. laidlawii* strain PG-8 in comparison to '*Ca.* P. asteris' strains OY-M and AY-WB (modified according to tRNAscan-SE analysis), '*Ca.* P. australiense' and '*Ca.* P. mali'. The green boxes indicate a conserved order of tRNAs in the acholeplasma and phytoplasma species. The green dotted lines indicate conserved synteny. The purple boxes highlight the conserved Val-Asn-tRNA organization in phytoplasmas, except for '*Ca.* P. australiense', while the purple dotted boxes mark the Asn-tRNA gene located next to one rRNA operon. Colors refer to the online version only.

Cell Division

The analyzed acholeplasma chromosomes encode the cell division protein FtsZ (Apal_08650, Abra_07640, ACL_0706) and the putative cell division protein SepF involved in septum formation (Apal_08680, Abra_07670, ACL_0703) that are not present in the completed phytoplasma genomes. The putative septation protein SpoVG is encoded in *A. palmae* (Apal_00960) and *A. laidlawii* (ACL_0017). SpoVG is essential for sporulation in *Bacillus megaterium* and *Bacillus subtilis* [Hudspeth and Vary, 1992], while its function in the genus *Acholeplasma* remains unclear.

Replication and Repair

In contrast to the phytoplasma genomes, those of the three acholeplasmas show a regular cumulative GC-skew

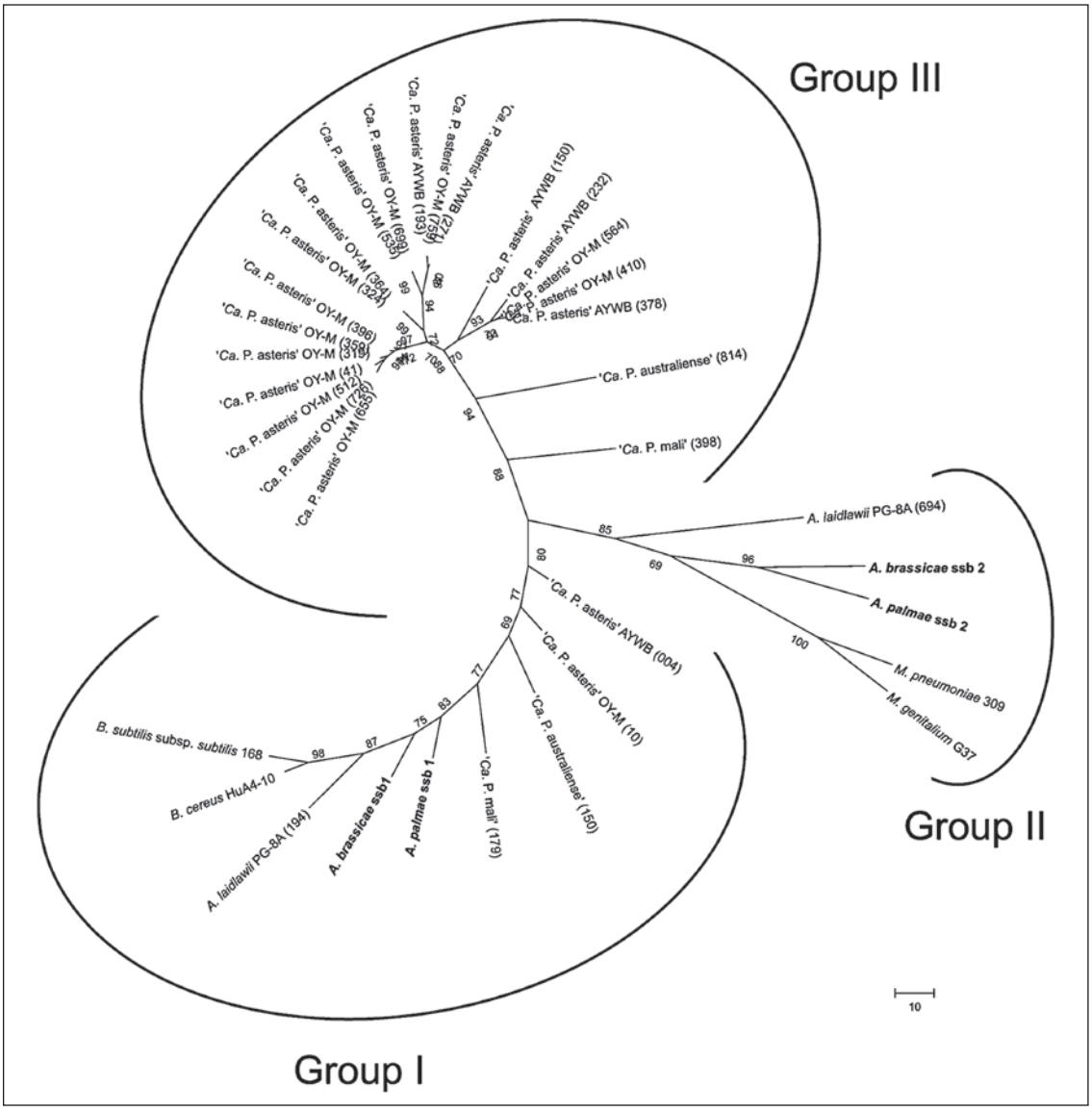


Fig. 3. Phylogenetic tree constructed by parsimony analyses of predicted *ssb* gene amino acid sequences of *A. palmae* and *A. brassicae*, obtained in this work (in bold), *A. laidlawii*, '*Ca.* P. mali', '*Ca.* P. australiense', '*Ca.* P. asteris', strains OY-M and AY-WB, *M. pneumoniae*, *M. genitalium*, *B. cereus,* and *B. subtilis* subsp. *subtilis*. Locus tag numbers are given in parentheses. Numbers on the branches are bootstrap values obtained for 1,000 replicates (only values above 70% are shown). The tree is drawn to scale, with branch lengths calculated using the average pathway method, and are in the units of the number of changes over the whole sequence. The scale bar represents 10 nucleotide substitutions.

pattern [(G-C)/(G+C)], which is associated with a characteristic switch in the coding strand preference (online suppl. material 6). Furthermore, they share a conserved *oriC* region containing the consensus DnaA-box motif 5′-TTATCCACA-3′ as determined for *A. laidlawii* [Lazarev et al., 2011] and genes comprising *dnaA*, *dnaN*, *yaaA*, *recF* and *gyrBA* downstream of this motif. The disruption of *dnaN* by a transposase is a particularity in *A. laidlawii*.

The local cumulative GC skew maximum indicating the predicted *terC* is located at position ±971 kb for *A. brassicae*, ±983 kb for *A. palmae* and ±716 kbp for *A. laidlawii,* and thereby shifted downstream. This may in-

dicate a disproportional effect during the bidirectional replication. All three chromosomes show a similar gene content at the suggested *terC* region comprising *deoB*, *xerD* and *hemN*. These genes have not been identified in the phytoplasma genomes.

A. palmae, A. brassicae and *A. laidlawii* encode the complete gene set and accessory genes necessary for the complex protein machinery at the replication fork *(A. brassicae*, *A. palmae: dnaABCDEGNQX*, *holA*, *ssb1/2*, *ligA*, *gyrAB*, *holB*, *polC*, *priA*, *topA*, *polA*, *parE*, *parC)*. Due to mobile genetic elements in phytoplasmas, many replication proteins are encoded several times in each genome. In the examined acholeplasmas, this situation is limited to the single-stranded binding protein *(ssb)*, which is encoded twice within these genomes and occurs as a multicopy gene in some phytoplasmas (online suppl. material 7). The *ssb* genes can be assigned to three phylogenetic clusters characterized by one conserved gene cluster shared by *Acholeplasma* and '*Ca*. Phytoplasma' (fig. 3: *ssb1,* group I) and two *ssb* gene clusters present in each genus and indicating a different evolutionary origin of these groups. It is remarkable that, in contrast to *ssb1*, transcripts of *ssb2* could not be identified in *A. palmae, A. brassicae, A. laidlawii* and '*Ca*. Phytoplasma mali' (online suppl. material 8). It is possible that this gene is not expressed, expressed at a low level or is differentially expressed, preventing detection of the transcripts.

In contrast to the phytoplasmas, the acholeplasmas possess a predicted type IV error-prone DNA polymerase (DinB), providing higher fitness by ensuring efficient replication [Yeiser et al., 2002]. A more comprehensive genetic content is also visible for the *rec*-dependent repair pathway encoded in the acholeplasmas. The chromosome of *A. brassicae* encodes the complete gene set (*recFGOANRDXU*) including the ATP-dependent helicase RecQ in contrast to the others (online suppl. material 9). The regulator RecX was not identified in *A. laidlawii*. Of the completely determined phytoplasmas, only '*Ca*. P. mali' encodes a minimal *rec*-dependent repair pathway.

All three acholeplasma species possess the genetic repertoire necessary to resolve the Holliday junction (RuvAB), the nucleotide excision repair system (UvrABCD) and the base excision repair system (MutTLSYM). However, *A. laidlawii* lacks the DNA glycosylase MutM. The *mutM* and *mutT* genes were also identified in all four phytoplasmas.

A. brassicae encodes the nuclease SbcCD, which harbors ATP-dependent double-strand DNA exonuclease activity and ATP-independent single-strand DNA endonuclease activity [Connelly and Leach, 1996]. The SbcCD protein complex contributes to the elimination or repair of DNA secondary structures and is thereby involved in the prevention of the propagation of palindromic phage sequences [Leach, 1994].

Nucleotide Metabolism

The three acholeplasma genomes encode the capacity to generate dNTPs from adenosine, guanosine, uracil and thymine (online suppl. material 10).

The pathway for the synthesis of dGTP and dATP from inosine monophosphate (IMP) is encoded in *A. palmae* and *A. brassicae*. Coding of initial proteins (GuAB) forming GMP from IMP were not identified for *A. laidlawii*. Other differences between the species are present in the ribonucleoside-diphosphate reductases that catalyze the conversion of nucleotides to deoxynucleotides. *A. palmae* and *A. laidlawii* are capable of this reaction by means of predicted ribonucleoside-diphosphate reductase class Ib (NrdEF), while *A. brassicae* only encodes the alpha chain (NrdE). *A. brassicae* encodes the subunits of class Ia ribonucleoside-diphosphate reductase (NrdAB) in contrast to the others. Nucleoside diphosphate kinase (Ndk), which performs the phosphorylation of GDP to GTP, was not predicted in the acholeplasma genomes. This feature is shared with other Mollicutes. It has been suggested that adenylate kinase (Adk) performs this function in phytoplasmas [Kube et al., 2012]. This assumption is supported by the class III ribonucleoside-triphosphate reductase (NrdD) and its activating protein (NrdG) predicted from all three genomes. NrdD reductase is inactivated by oxygen in *Escherichia coli* [Garriga et al., 1996], and thereby might represent an adaptation to low oxygen environments. The glutaredoxin-like protein NrdH acting as the putative functional electron donor for the NrdEF ribonucleotide reductase and the ribonucleotide reduction simulating protein NrdI [Gon et al., 2006] were not predicted for *A. brassicae* in contrast to the other two acholeplasmas.

IMP, however, has to be provided to the cellular metabolism. The genomes of *A. palmae and A. brassicae* encode a permease that may import xanthine or hypoxanthine, which is not predicted for *A. laidlawii*. IMP may be generated from hypoxanthine and 5-phospho-α-D-ribose-1-diphosphate mediated by purine nucleoside phosphorylase (DeoD) and 5′-nucleotidase (SurE), or directly by hypoxanthine-guanine phosphoribosyltransferase (HPT). The genes for these steps are present in the three acholeplasmas.

Folate

Bacteria depend on tetrahydrofolates as important cofactors in one-carbon transfer reactions such as the synthesis of formylmethionyl-tRNA [de Crecy-Lagard et al., 2007]. Elements for the generation of the folate pool were identified in *A. laidlawii* [Lazarev et al., 2011] and the other two acholeplasmas share key elements of the folate salvage pathway (online suppl. material 11). The dihydropteroate synthase (FolP) performing the initial step of 7,8-dihydropterate was not identified, like in many other bacteria lacking the genetic repertoire for the de novo synthesis and dependent on the uptake from the environment. The following steps resulting in the conversion to tetrahydrofolate and subsequent 5,10-methylenetetrahydrofolate are encoded in all three acholeplasma genomes. In contrast, phytoplasmas cannot convert intermediates of the folate pool. They encode only parts of the involved pathways and even these components may represent pseudogenes as shown for the '*Ca.* P. asteris'-related strain CPh [Davis et al., 2005].

Translocation Processes at the Cytoplasmic Membrane

The three acholeplasma species share a rich repertoire of ABC transporters (fig. 4, online suppl. material 12, 13) enabling the ATP-driven import of methionine, amino acids/oligopeptides, spermidine/putrescine (PotABC), carbohydrates/sugars, phosphate (PstABC), cobalt (CbiOQ) and metal ions (ZnuABC). The ABC transporter for branched-chain amino acids is encoded in *A. brassicae* and *A. laidlawii*, while the substrate-binding protein was not predicted for *A. palmae*. Deduced from the genomes, amino acid/polyamine transporter exchanges a standard α-amino acid for a polyamine in the cells of *A. brassicae* and *A. palmae*.

The highest number of subunits assigned to the uptake of carbohydrates was identified in *A. brassicae*. Glucose, fructose and galactose have been described as entry molecules for *A. laidlawii* [Lazarev et al., 2011]. It is remarkable that *A. brassicae* and *A. laidlawii* also possess a predicted C4-dicarboxylate transporter (DcuC).

A. laidlawii and *A. palmae* lack the Fe^{2+} importer (FeoAB), which is predicted for *A. brassicae*. An antimicrobial peptide transporter was predicted for *A. palmae* and *A. laidlawii*, while a bacteriocin exporter was predicted only for *A. palmae*. Furthermore, major facilitator superfamily transporters for general transport of small solutes were predicted in all three *Acholeplasma* species.

Uniporters for formate/nitrite (FNT) and the proton-coupled import of thiamine (YuaJ) are encoded in the genomes of all three acholeplasmas. *A. palmae* and *A. laidlawii* contain uniporters (TrkA) and symporters (TrkG) for potassium uptake. *A. brassicae* performs this uptake by the uniporters KtrC and KtrD. A symporter for the glutamate (GltS) encoding gene is shared by all three genomes. Unique symporters are GntP (H^+/gluconate) in *A. palmae* and MelB (Na^+/melibiose) in *A. brassicae*.

The gene content for antiporters remains heterogeneous. The Na^+/H^+ antiporter (NhaB) is predicted for *A. laidlawii* and *A. brassicae*, while the K^+/H^+ antiporter (KefB) is apparently present only in *A. laidlawii*. The gene for a Na^+/Ca^{2+} antiporter (NCX) is shared by the three acholeplasmas.

A rich repertoire of p-type ATPases is shared for the export of H^+ (HTPase) and Cu^{2+} (CopA), as well as for ATPases mediating the import of metal ions such as Mg^{2+}, Cd^{2+}, Cu^{2+}, Zn^{2+}, Ca^{2+} and Na^+. In addition, the *A. brassicae* and *A. laidlawii* genomes encode a Cd^{2+}-translocating p-type ATPase (CadA). These ATPases were not predicted in phytoplasma genomes.

The complete gene set for the F_0F_1-type ATP synthase (AtpABCDEFGH) was identified in the three acholeplasma genomes, as well as that for a V-type ATP synthase (NtpABCDEFGKI) for the release of Na^+. Similarly, the uptake of Na^+ is mediated by the multidrug efflux pump (NorM) is predicted for the three *Acholeplasma* species that have been studied.

A. brassicae encodes the complete units forming the NADH oxidoreductase complex (RnfABCDGE), which might be incomplete in *A. palmae* (RnfACDE) and *A. laidlawii* (RnfCDE). *A. brassicae* and *A. laidlawii* also differ from *A. palmae* by encoding the Na^+-transporting ubiquinone oxidoreductase complex (NqrABCDEF).

The uptake of Mg^{2+} is performed by the transporter MgtE in all three species. In addition, *A. brassicae* and *A. laidlawii* encode a transporter for the import of Cr^{2+} (ChrA). In *A. brassicae,* Fe^{2+} ions can be taken up by the permease FeoAB and other divalent inorganic ions such as Mg^{2+} and Co^{2+} by CorA. One may speculate whether homeostasis of the divalent ion pool is regulated by the Co^{2+}, Zn^{2+} and Cd^{2+} efflux protein (RND). Moreover, *A. palmae* and *A. laidlawii* possess a voltage-gated Cl^- channel protein. These channels are activated by voltage to import Cl^- ions. The major intrinsic proteins import small neutral solutes in *A. brassicae* and *A. laidlawii*.

Apparently, the three acholeplasmas can handle mechanical stress by the small (MscS) and large conductance mechanosensitive channel proteins (MscL).

A surprising feature of the plant-derived isolates *A. brassicae* and *A. palmae* is the prediction of a putative auxin efflux carrier protein for exporting indole-3-ace-

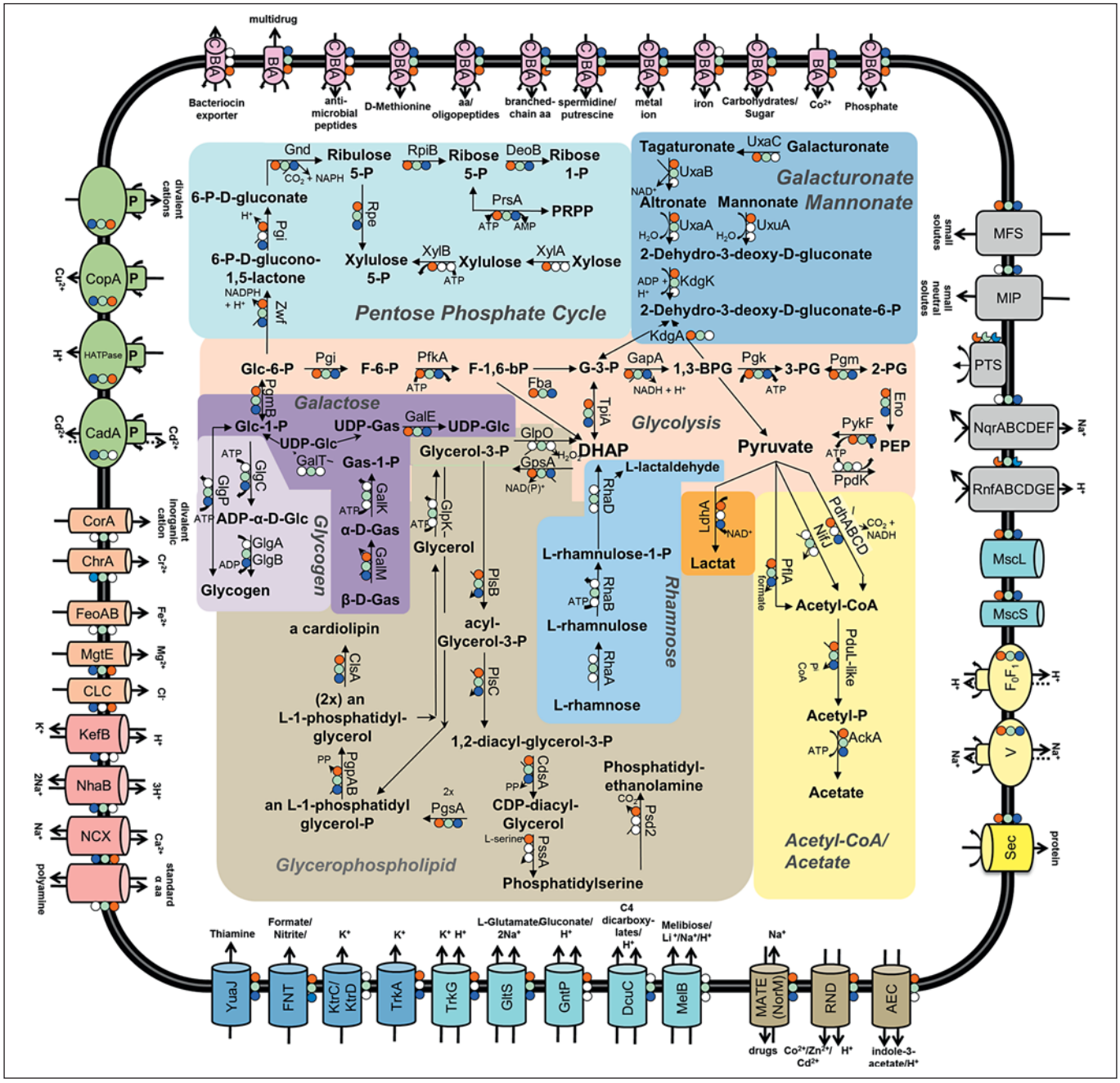


Fig. 4. Genetic elements involved in transport at the membrane and key elements of the carbohydrate metabolism in *A. palmae* (red dot), *A. brassicae* (green dot) and *A. laidlawii* (blue dot). Genes not predicted for an organism were indicated by a white dot. Curved arrows indicate ATP hydrolysis-driven actions.

tate, which represents an important growth factor in plants. A functional system could manipulate the plant growth. The benefit for a saprophytic bacterium remains unclear. Possibly, dissemination of bacteria by insects is improved or acholeplasmas also take part in degradation of living tissue. The last scenario suggests a potential phytopathogenicity that has been shown in recent laboratory experiments with *Oryza sativa* and the production of virulence-associated extracellular membrane vesicles by *A. laidlawii* [Chernov et al., 2012].

All three *Acholeplasma* species are predicted to harbor components of the phosphotransferase system (PTS). The PTS consists at least of three main components: the membrane integral permease enzyme II and the two cytoplasmic components enzyme I and HPr protein [Deutscher et al., 2006]. The sugar-specific enzyme II phosphorylates and transports specific carbohydrates across the membrane. The components PtsI, PtsP and PtsA belong to the cytoplasmic enzyme complex EI. PtsI encodes for the sugar transport and phosphorylation [Tchieu et al., 2001], whereas PtsP plays a role in the nitrogen-metabolic PTS [Rabus et al., 1999] and PtsA is a homolog of FryA, which is responsible for the fructose transport (EC2.7.1.69). Hpr proteins are phosphocarriers: Hpr1, a histidine phosphorylation site protein; Hpr2, a serine phosphorylation site protein, and HprK, a kinase/phosphorylase protein. All three *Acholeplasma* species possess Hpr1, Hpr2 and HprK proteins. However, each species encodes a different EI enzyme: PtsP in *A. brassicae*, PtsI in *A. palmae* and PtsA in *A. laidlawii*. The distribution of EII proteins also seems to vary between the three species. *A. palmae* harbors two fructose-transporter fragments of EII and one glucose-specific EII component. In contrast, the *A. laidlawii* genome encodes only one fructose fragment of EII and one glucose-specific EII component. *A. brassicae* seems to be devoid of an EII component. Hoischen et al. [1993] have postulated that *A. laidlawii* has an incomplete PTS since they detected enzyme activities of EI and HPr, but not of EII. While *A. palmae* may contain a functional PTS, the situation remains unclear for *A. brassicae*. No single component of an EII complex could be predicted. Phytoplasmas lack this genetic repertoire, but they share a poorly conserved cohort of ABC-transporters [Kube et al., 2012], which is also encoded in the three acholeplasma genomes (ABC transporters for carbohydrates/sugars, spermidine/putrescine, oligopeptides/amino acids, D-methionine, metal ions and cobalt). This is also the case for a multidrug transporter, a Na^+-driven multidrug efflux pump and a cation transporting p-type ATPase. Phytoplasma genomes differ in that they encode a symporter for Na^+/malate or citrate.

Secretion

Acholeplasmas and phytoplasmas share central components of the *Sec*-dependent secretion system, except for the preprotein translocase SecG, which was only predicted for the three *Acholeplasma* species (online suppl. material 14). SecG is not essential for the function of the system in general, but it might be of importance if secretion should be carried out at temperatures below 20°C as it was shown for *E. coli* [Nishiyama et al., 1994]. The examined *Acholeplasma* species are frequently exposed to temperature shifts in contrast to the intracellularly localized phytoplasmas. However, it should be also considered that a different protein could substitute the function of *Sec*-channel formation in conjunction with the integral membrane proteins SecY and SecE.

The shares of predicted proteins secreted by the *Sec*-dependent system are 11% for *A. brassicae* (190 proteins), 10% for *A. palmae* (137 proteins) and 8% for *A. laidlawii* (110 proteins). They are remarkably higher than the 4.8–5.4% predicted for the phytoplasmas (online suppl. material 15, 16).

According to the prediction by Phobius, all three *Acholeplasma* species secrete some proteins that would not have been expected to occur in the extracellular space. These include a prephenate dehydrogenase (TyrA), which forms 4-hydroxyphenylpyruvate from prephenate and NAD^+ within the tyrosine biosynthesis. Only *A. palmae* and *A. brassicae* are also predicted to secrete glucose-6-phosphate-1-dehydrogenase, an enzyme constituent of the pentose phosphate pathway. *A. laidlawii* and *A. palmae* are predicted to secrete L-lactate dehydrogenase. *A. palmae* is the only one predicted to secrete acetoin dehydrogenase (BudC), mediating the conversion of pyruvate to 2,3-butanediol, and *A. brassicae* is the only one possibly secreting an UDP-glucose/GDP-mannose dehydrogenase. *A. laidlawii* is predicted to secrete a NAD (FAD)-dependent dehydrogenase and a phytoene dehydrogenase. In contrast, '*Ca*. P. asteris' strain AY-WB and '*Ca*. P. mali' are the only phytoplasma species which are predicted to secrete a glycerol-3-phosphate dehydrogenase. The other phytoplasma species were not predicted to secrete any dehydrogenase.

A. brassicae may export ribose-phosphate pyrophosphokinase, which would allow the extracellular ATP-dependent formation of AMP. Furthermore, all three acholeplasmas seem to export ATP-dependent uridine kinase forming UMP. In contrast, the four phytoplasma genomes may secrete guanylate kinase for the formation of GDP. '*Ca*. P. mali' is predicted to export adenylate kinase for the formation of AMP. Furthermore, the three acholeplasmas probably secrete endonucleases and ribonucleases for cleavage of DNA and RNA.

Further potentially secreted proteins include 6-phosphofructokinase (PfkA), which could catalyze the extracellular formation of fructose 1,6-bisphosphate. Such a scenario would be remarkable because the cell would not provide the needed ATP in this reaction. Notably, each

acholeplasma genome contains two 6-phosphofructokinase-encoding genes. Only one deduced protein per genome encodes a signal peptide (Abra_12780, Apal_11830, ACL_1092) clustered together on the chromosome with the other copy without a signal peptide (Abra_12770, Apal_11800, ACL_1091). However, the uptake of such an extracellularly phosphorylated hexose remains unclear. Contrasting the acholeplasmas, phytoplasma genomes encode only one copy of PfkA without a signal peptide. Expectedly, these are more closely related to PfkAs in the acholeplasmas, and also have no signal peptide (data not shown). An uptake of fructose 1,6-bisphosphate from the environment would compensate an absent PTS system in phytoplasmas. Direct access to a phosphorylated hexose might be possible for these intracellular parasites [Kube et al., 2012].

A. laidlawii and *A. brassicae* seem to export α-amylase, which is needed to break down starch or glycogen to glucose and maltose. Both plant-derived *Acholeplasma* species might secrete a pectate lyase, which cleaves galacturonans to gain oligosaccharides. This feature may be interpreted with respect to the saprophytic lifestyle as the exopolygalacturonate lyase of *A. brassicae* and the pectin methylesterase of *A. palmae* are also predicted.

Proteins predicted to be secreted by phytoplasmas were not encoded in the acholeplasma genomes. This may indicate an evolutionary specialization of the phytoplasmas and might be due to a different origin of these genes that also include encoded effector proteins. Results obtained by Phobius analysis include experimentally verified effectors of phytoplasmas, such as *tengu* of the '*Ca*. P. asteris' strain OY-M [Hoshi et al., 2009], which is responsible for dwarfism in plants, and SAP42, SAP11 and SAP30 of the '*Ca*. P. asteris' strain AY-WB [Bai et al., 2009]. In addition, higher percentages of posttranslationally processed proteins carrying a signal peptide and transmembrane regions were observed in *A. brassicae* (3%) and *A. palmae* (2.2%) in comparison to *A. laidlawii* (1.5%) and the four phytoplasmas (1.2–1.6%). These proteins are embedded in the membrane and are of importance for the phytoplasma host/vector interaction [Hogenhout et al., 2008]. No homolog of prominent immunodominant membrane proteins described for phytoplasmas were identified in the acholeplasma genomic data. Another important and suggested virulence-related protein in phytoplasmas is encoded by *hflB* and encoded by *A. brassicae* (Abra_16360), *A. palmae* (Apal_13970) and *A. laidlawii* (ACL_1386). HflB is a Zn-dependent protease, which is predicted to be secreted by '*Ca*. P. mali' and analyzed in detail in severe and mild strains [Seemüller et al., 2013]. In addition, an AAA+ ATPase was identified in the genome of *A. brassicae* (Abra_01290). However, the proteases are predicted to be membrane bound, while the AAA+ ATPase is limited to the cytosol.

Besides the shared *Sec*-dependent secretion pathway, genes predicted to encode IcmE/DotG of the core complex of the type IVB secretion system (T4BSS) were identified in '*Ca*. P. asteris' strains AY-WB [Bai et al., 2006] and OY-M, '*Ca*. P. australiense' [Tran-Nguyen et al., 2008], '*Ca*. P. mali' (YP_0020041271), and in several draft genomes of the phytoplasma 16SrIII group [Saccardo et al., 2012]. T4BSS might be also involved in the release of effector proteins [Nagai et al., 2002]. IcmE/DotG plays a central role in the core complex of the T4BSS of Gram-negative *Legionella* spp. [Nagai and Kubori, 2011]. However, no component of the T4BSS was identified in the three *Acholeplasma* species.

Carbohydrate Metabolism

A. laidlawii [Lazarev et al., 2011], *A. palmae* and *A. brassicae* have a considerably more complex carbohydrate metabolism (fig. 4) than phytoplasmas. While the acholeplasmas differ mainly in particular bypaths, phytoplasmas do not encode complete pathways. Detailed differences are presented in the following subsections.

Glycolysis

Like for most fermenting Mollicutes, glycolysis represents the central pathway to generate ATP. It has been reconstructed for *A. laidlawii* [Lazarev et al., 2011] and key elements were identified in this study for *A. brassicae* and *A. palmae*. Glycolysis from D-glucose-6-phosphate represents the central ATP providing pathway (online suppl. material 17). Besides the above-mentioned phosphofructokinase (PfkA), pyruvate kinase (PykF) is encoded twice in *A. brassicae*. This species also encodes the genetic repertoire for the generation of phosphoenolpyruvate from pyruvate by pyruvate phosphate dikinase (PpdK), which was not identified in the other genomes.

Furthermore, *A. brassicae* also possesses the gene sets to generate dihydroxyacetone-phosphate from glycerol (GlpK, GlpA/GlpO) [Bischof et al., 2009; Fujita, 2009] including the antiterminator protein (GlpP). Glycerol kinase (GlpK) is encoded twice in the genome (online suppl. material 18). This pathway may result in the release of toxic H_2O_2, which is a known virulence factor in *Mycoplasma mycoides* [Bischof et al., 2009; Vilei and Frey, 2001]. However, no transporter subunit for glycerol uptake (GtsABC) was identified.

Pyruvate and Acetyl-CoA Metabolism

The gene content associated with the metabolism of pyruvate differs between the three *Acholeplasma* species (online suppl. material 19). Pyruvate is connected to glycolysis, lactate and actetyl-CoA metabolism in all three acholeplasmas. *A. palmae* and *A. laidlawii* encode a lactate dehydrogenase (LdhA), which enables them to reduce pyruvate to (*R*)-lactate. Phytoplasmas lack the *ldhA* gene.

The pyruvate dehydrogenase multienzyme (PdhABCD) complex is encoded in *A. palmae* and *A. laidlawii,* resulting in the formation of acetyl-CoA. Alternatively, acetyl-CoA and formate can be generated from pyruvate and CoA by the pyruvate formate lyase (PflA) in all three acholeplasmas. Only *A. brassicae* encodes a pyruvate-flavodoxin oxidoreductase (NifJ), offering an alternative way for the conversion.

In contrast to *Mycoplasma* species, the three acholeplasmas lack the phosphotransacetylase (Pta), allowing the formation of acetyl-phosphate. The lack of Pta is also shared by the phytoplasmas [Kube et al., 2012]. Phytoplasmas and the analyzed acholeplasmas encode acetate kinase (AckA), converting acetyl-phosphate to acetate and ATP. Thereby, a third ATP-generating reaction can be used in addition to the two respective reactions of the glycolysis. Formation of the required acetyl-phosphate could be achieved by propanoyltransferase PduL (PduL-like transferase) (fig . 4), as recently suggested for the phytoplasmas [Kube et al., 2012]. However, acetyl-CoA may also be provided from pyruvate generated from imported malate (symporter MleP/CitS) and converted by the conserved malate dehydrogenase (SfcA) in phytoplasmas [Kube et al., 2012]. Neither protein was identified in the acholeplasmas.

Glycerophospholipid Metabolism

Dihydroxyacetone phosphate can be converted by all three acholeplasmas to major phospholipids such as phosphatidylglycerol and cardiolipin. Released glycerine during the synthesis of cardiolipin can be phosphorylated by glycerol kinase (GlpK), resulting in glycerol-3-phosphate in *A. brassicae*. *A. palmae* differs from *A. brassicae* and *A. laidlawii* by the presence of phosphatidylserine synthase (PssA) and phosphatidylserine decarboxylase (Psd), thus providing the enzyme machinery to generate phosphatidylethanolamine (online suppl. material 20).

Pentose Phosphate Pathway

The three acholeplasma genomes encode key elements of the pentose phosphate cycle (PPC) for the interconversion of the sugar phosphates and provision of NADPH (online suppl. material 21). The complete oxidative-branch of the PPC is encoded in *A. palmae* and *A. laidlawii,* and includes glucose-6-phosphate-1-dehydrogenase (Zwf), 6-phosphogluconolactonase (Pgi) and 6-phosphogluconate dehydrogenase (Gnd). In contrast, Pgi was not predicted for *A. brassicae*. Furthermore, all three acholeplasma genomes encode ribulose-5-phosphate-3-epimerase (Rpe), ribose-5-phosphate isomerase B (RpiB) and the transketolase (TktA), resulting in the formation of D-glyceraldehyde-3-phosphate, which may enter glycolysis. Intermediates of the PPC such as D-ribose-5-phosphate and D-xylulose-5-phosphate are connected to a by-pass. D-ribose-5-phosphate may be generated by all three acholeplasmas from 5-phosphoribosyl 1-pyrophosphate (PRPP) and AMP under formation of ATP, and on the other hand mediated by a phosphopentomutase (DeoB) from D-ribose-1-phosphate.

A. palmae is predicted to transform D-xylose via another pathway to D-xylulose-5-phosphate. D-xylose may be converted by the xylose-isomerase (XylA) to D-xylulose and phosphorylated by the xylulokinase (XylB). The uptake and the utilization of D-xylose, which is also known as wood sugar, may be interpreted as a link to this plant-derived isolate. Neither gene was identified in the genomes of *A. brassicae* and *A. laidlawii*. The PPC is absent in the four phytoplasma genomes.

Galacturonate Metabolism

In contrast to *A. laidlawii*, *A. palmae* and *A. brassicae* are predicted to be capable of galacturonate catabolism. They would access D-galacturonate as a carbon source by the enzyme sequence of D-galacturonate isomerase (UxaC), altronate oxidoreductase (UxaB), D-altronate dehydratase (UxaA) and 2-keto-3-deoxygluconokinase (KdgK). 2-dehydro-3-deoxy-gluconate-6-phosphate can be coverted to D-glyceraldehyde-3-phosphate and pyruvate by the 2-keto-3-deoxygluconate-6-phosphate aldolase (KdgA). This pathway is absent in *A. ladlawii* (online suppl. material 22). *A. palmae* differs from *A. brassicae* by encoding a D-mannonate dehydratase (UxuA) allowing the degradation of D-mannonate. However, D-mannonate oxidoreductase (UxuB) and D-glucuronate isomerase (UxuC) could not be predicted from the genomes of *A. palmae* and *A. brassicae*. The genetic repertoire for the galacturonate metabolism is missing in phytoplasmas except for the multifunctional enzyme 2-keto-3-deoxygluconate-6-phosphate aldolase (Eda) in '*Ca*. P. mali'.

Galactose Degradation

β-D-galactose is suggested to be taken up by *A. brassicae*. This hexose is phosphorylated and converted by the Leloir pathway to UDP-D-glucose and α-D-glucose-1-phosphate, which allows utilization as β-D-glucose-6-phosphate in the glycolysis. For the required pathway, *A. brassicae* is predicted to harbor galactose-1-epimerase (GalM), galactokinase (GalK), galactose-1-phosphate uridylyltransferase (GalT), UDP-glucose 4-epimerase (GalE) and UDP-glucose-pyrophosphorylase (GalU). It was not possible to identify *galK* and *galT* in the genomes of *A. palmae* and *A. laidlawii*, while the GalE and GalU are encoded in all three genomes providing UDP-D-glucose and UDP-D-galactose (online suppl. material 23).

Glycogen Metabolism

α-D-glucose-1-phosphate derived from the galactose metabolism and ATP may enter glycogen biosynthesis via predicted glucose-1-phosphate adenylyltransferase (GlgC), glycogen synthase (GlgA) and the 1,4-α-glucan branching enzyme (GlgB) in *A. brassicae* and *A. laidlawii,* but not in *A. palmae* which lacks this pathway. Glycogen can be remobilized by the glycogen phosphorylase (GlgP) releasing α-D-glucose-1-phosphate. However, the genetic repertoire needed for efficient degradation was not predicted (online suppl. material 24).

Rhamnose Metabolism

The genetic repertoire for L-rhamnose degradation comprising L-rhamnose isomerase (RhaA), L-rhamnulose kinase (RhaB) and rhamnulose-1-phosphate aldolase (RhaD) was only predicted for *A. brassicae*. A transcription regulator and genes encoding a carbohydrate ABC-transporter flank the *rha*-operon in *A. brassicae,* indicating a functional unit. Dihydroxyacetone phosphate may enter the glycolysis, while the processing of the L-lactaldehyde remains unclear (online suppl. material 25). The genetic repertoire for utilization of rhamnose, galactose and glycogen could not be predicted for the phytoplasmas.

Fatty Acids

Acetyl-CoA acts as an entry molecule for palmitate biosynthesis in all three acholeplasmas, which share a common gene set. Malonyl-CoA is synthesized from acetyl-CoA, bicarbonate and ATP by acetyl-CoA carboxylase (AccBCDA) and is used for the generation of palmitate by an analogous enzyme cascade (FabGZIBZF) in all three acholeplasmas (online suppl. material 26). Palmitate may represent a constituent of phospholipids in the cytoplasmatic membrane of acholeplasmas. The pathway was not reconstructed in phytoplasmas that seem to depend on their hosts in this respect.

In contrast to *A. laidlawii* and *A. palmae*, *A. brassicae* possesses the genetic repertoire to generate butanoyl-CoA from short-chained butanoate by phosphate butyryltransferase (Abra_11920; EC 2.3.1.19) and butyrate kinase (Abra_11910; EC 2.7.2.7). *A. brassicae* also encodes 4-hydroxybutyrate CoA transferase (Abra_09260; EC 2.8.3.8), which may catalyze the formation of 4-hydroxybutyryl-CoA from 4-hydroxybutyrate and acetyl-CoA as described for *Clostridium aminobutyricum* [Scherf and Buckel, 1991].

Isoprenoid Synthesis

Terpenoids are generated by the mevalonic acid pathway in *A. palmae* and *A. laidlawii* and by the 2-C-methyl-D-erythritol-4-phosphate/1-deoxy-D-xylulose-5-phosphate pathway (MEP/DOXP pathway) in *A. brassicae*. Both pathways generate isopentyl diphosphate (IPP). The mevalonic acid pathway starts from acetyl-CoA and the MEP/DOXP pathway from D-glyceraldehyde-3-phosphate. The MEP/DOXP pathway in *A. brassicae* contains a putative bifunctional protein 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase/2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (IspD/IspF), which is also described, e.g. strain HTCC2255 (Acc. No. Q0FAI8; Rhodobacterales). In the final steps, the intermediate 2-C-methyl-D-erythritol-2,4-cyclodiphosphate is reduced, resulting in the formation of 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate, which is converted by isopentenyl diphosphate isomerase to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate dimethylallyl diphosphate and isopentenyl diphosphate. This isomerase was not predicted for *A. brassicae,* but for *A. palmae* and *A. laidlawii*. The latter two are restricted to the mevalonic acid pathway, providing only isopentenyl diphosphate.

1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate dimethylallyl diphosphate and isopentenyl diphosphate may enter the trans-lycopene biosynthesis pathway providing *all-trans*-ζ-carotene to the tree acholeplasmas. However, since geranyltranstransferase was only identified in *A. laidlawii,* it remains unclear if this pathway still is functional in *A. palmae* and *A. brassicae* (online suppl. material 27). Phytoplasmas are lacking this pathway.

Amino Acid Metabolism

Pathway fragments for the synthesis or conversion of amino acids are encoded in the three acholeplasma genomes. For example, fumarate can be converted,

mediated by adenylosuccinate lyase (PurB: Apal_03580, Abra_01980; ACL_1332) and synthase (PurA: Apal_03600, Abra_00280) to L-aspartate. Encoding of PurB in all three genomes may be linked to the de novo biosynthesis of purine nucleotides. However, the majority of genetic modules indicates the conversion of intermediates and is connected to ABC transporter genes for the uptake of oligopeptides and amino acids. L-serine, for example, may not be generated by *A. palmae* and *A. brassicae,* but both species possess a L-serine deaminase (Apal_05700; Abra_12440), allowing degradation and usage of L-serine as carbon-source.

In principle, D-erythrose-4-phosphate can be generated via several pathways. *A. palmae* forms it using the fructose-6-phosphate phosphoketolase (Xfp), which is probably not the case for the two other acholeplasmas. In contrast, the enzymatic repertoire for processing D-erythrose-4-phosphate and phosphoenolpyruvate to chorismate, and subsequently to prephenate, is shared by all three acholeplasmas (online suppl. material 28). Prephenate is suggested to be processed by a prephenate dehydratase (PheA) and the NAD-dependent prephenate dehydrogenase (TyrA). A candidate for an aminotransferase of aromatic amino acids necessary to gain L-phenylalanine and L-tyrosine was not identified in the three acholeplasmas. However, it is likely that such a transferase is encoded. The aminotransferase Ilve, encoded in *A. laidlawii*, was not identified in the two other acholeplasmas.

Conclusions

Comparative analysis of the genomes of *A. brassicae* strain O502 and *A. palmae* strain J233 reported here, and the previously published genomes of *A. laidlawii* strain PG-8, '*Ca.* P. asteris' strains OY-M and AY-WB, '*Ca.* P. australiense', and '*Ca.* P. mali' highlight a shared restricted genetic repertoire for metabolism including glycolysis (incompletely encoded in '*Ca.* P. mali'), minimal set of membrane transporter systems and a dependency on the uptake of amino acids. Differences are obvious in genome content related to the colonization of environmental habitats by acholeplasmas and the dependency on the plant phloem and insect vectors by phytoplasmas. The absence of most pathways encoded in acholeplasma genomes in those of the phytoplasmas should mirror the lifestyle of the latter, i.e. colonization of nutrient-rich habitats such as the phloem sap and also the insect vectors [Kube et al., 2012]. Consequently, the examined acholeplasma genomes are separated from the phytoplasmas by a relative richness of genes for transporters and carbohydrate metabolism. Besides central glycolysis, they include metabolic pathways for pyruvate acetyl-CoA, pentose phosphate, galacturonate, galactose, glycogen and rhamnose. The F_0F_1-ATP synthase and the Rnf-complex may also be limited to the acholeplasmas. The genetic elements for the synthesis of isoprenoids, fatty acids and a cryptic amino acid metabolism are also characteristic of the acholeplasma genomes. These features may be interpreted with respect to the saprophytic lifestyle and the limited access to nutrients.

Conserved metabolic genes of phytoplasmas such as the malate dehydrogenase SfcA, which is central to a suggested alternative pathway that provides energy to phytoplasmas [Kube et al., 2012], were not identified in the three acholeplasma genomes. This might be due to the data basis which is still weak. In contrast, the putatively acetylphosphate-forming PduL-like enzyme is shared in both genera.

With respect to the relationship of both genera, it is noteworthy that the acholeplasmas are predicted to secrete metabolic key enzymes such as 6-phosphofructokinase, which may carry a signal peptide or not. It is conceivable that the extracellular utilization of metabolic intermediates was a key step on the way to intracellular parasitism of phytoplasmas and that this strategy is already encoded in parts in the analyzed acholeplasmas.

The integration and establishment of proteins involved in host interaction such as immunodominant proteins and effectors were not predicted for the acholeplasmas, indicating a deep split between the two genera. A rapid phytoplasma evolution resulting in a manipulation of plant and insect vectors driven by horizontal gene transfers and duplication events resulted in genome instability [Bai et al., 2006; Wei et al., 2008], which is not a pronounced feature of the three acholeplasma genomes.

Material and Methods

Bacterial Strains

A. brassicae strain O502 and *A. palmae* strain J233 were provided by Jerry K. Davis (Purdue University School of Veterinary Medicine, West Lafayette, Ind., USA) from the strain collection of the International Organization for Mycoplasmology. Single colonies were used to inoculate cultures of *A. brassicae* and *A. palmae,* grown at 30 and at 37°C, respectively, in SP4 medium [Tully et al., 1994].

DNA Extraction, Library Construction and Sequence Determination

Genomic DNA was isolated from liquid cultures using the Qiagen Genomic DNA Kit (Qiagen, Hilden, Germany) according to

the manufacturer's instructions. Shotgun libraries with different insert sizes were constructed for both strains. Clone libraries with 1.3- and 2.3-kb inserts and fosmids with 37 kb (Copy Control Fosmid Library Construction Kit; Epicentre, Madison, Wisc., USA) were generated as described [Kube et al., 2005]. Recombinant DNA was isolated by modified alkaline lysis (Automated Plasmid Isolation System; Max Planck Institute for Molecular Genetics, Berlin, Germany) followed by cycle sequencing (BigDye Terminator V3.1; Applied Biosystems, Darmstadt, Germany) and end sequences determined on ABI 3730XL systems (Applied Biosystems). In addition, clone-independent shotgun sequences were obtained by pyrosequencing using the GS FLX system (454 Life Sciences/Roche, Branford, Conn., USA) according to the manufacturer's instructions.

Data obtained by the different approaches were assembled by the Celera assembler version 4.4 [Myers et al., 2000]. Data were edited in Consed [Gordon et al., 1998] and finishing data were obtained from primer walking on clones and long-range PCR products.

Data Analysis

Complete genome sequences were analyzed using Glimmer [Delcher et al., 1999], HTGA [Rabus et al., 2002] and Artemis [Rutherford et al., 2000]. The software packages tRNAscan-SE [Lowe and Eddy, 1997] and RNAmmer [Lagesen et al., 2007] were used for tRNA and rRNA prediction. InterPro [Hunter et al., 2012], KEGG (www.genome.jp/kegg/) and MetaCyc [Caspi et al., 2012] were included in pathway reconstruction. The number of paralogs was estimated using BlastClust (www.ncbi.nlm.nih.gov/). Deduced protein sequences from the chromosomes including data of potential pseudogenes were compared applying a coverage minimal length of 70%. Annotated pseudogenes were incorporated in this analysis because of their influence on genome size and the lack of experimental confirmation of their assignment. The percentage of paralogs of the total number of potential protein-encoding genes includes the annotated pseudogenes as a consequence. Shared proteins were calculated by reciprocal BLASTP [Kube et al., 2010] using *A. brassicae* and *A. palmae* as subjects and applying cutoff values of a minimal identity of 30%, a minimal alignment coverage of 70% and an e-value of 1e-10. The applied pairwise reciprocal comparison discriminates the proteins encoded by multicopy genes in phytoplasmas and thereby should provide a better overview of the shared informational content.

The software SIGI-HMM [Waack et al., 2006] was used to identify genes showing unusual codon usage or horizontal gene transfers in the genomes of *A. brassicae*, *A. palmae*, *A. laidlawii*, '*Ca*. P. australiense', '*Ca*. P. mali' and the '*Ca*. P. asteris' strains OY-M and AY-WB. Standard parameters were used, but the sensitivity value was changed to 0.8. Three misleading overpredicted genes located in the rRNA operon of '*Ca*. P. asteris' strains (locus tags: AYWB_261, PAM_250, PAM_492) were excluded from the codon usage analysis.

Phobius [Kall et al., 2007] was used for the prediction of proteins carrying transmembrane regions and/or signal peptides in the deduced proteins of *A. brassicae*, *A. palmae*, *A. laidlawii*, '*Ca*. P. australiense', '*Ca*. P. asteris' strains OY-M and AY-WB, and '*Ca*. P. mali'.

Phylogenetic Analysis of ssb Genes

Predicted amino acid sequences of both annotated *ssb* gene copies of *A. laidlawii* strain PG-8A (NC_010163.1), '*Ca*. P. mali' (NC_011047.1), '*Ca*. P. australiense' (NC_010544.1), all annotated copies of the gene (15 and 5, respectively) in '*Ca*. P. asteris', strains OY-M (NC_005303.2) and AYWB (NC_007716.1), *Mycoplasma pneumoniae* strain 309 (NC_000908.2), *M. genitalium* strain G37 (NC_000908.2), *Bacillus cereus* strain HuA4–10 (ZP_17497163) and *B. subtilis* subsp. *subtilis* strain 168 (AL009126.3) were retrieved from GenBank (http://www.ncbi.nlm.nih.gov). Besides the Mollicutes, the closest relatives of acholeplasmas, two species from the *Bacillus* genus, were selected as the representative genus with highest similarity in the BLASTP program. The retrieved *ssb* gene sequences were aligned with those obtained in this work using CLUSTAL W [Thompson et al., 1997] from the Molecular Evolutionary Genetics Analysis program-MEGA5 [Tamura et al., 2011] and adjusted manually.

The evolutionary history was inferred by using the maximum parsimony method. The maximum parsimony tree was obtained using the subtree-pruning-regrafting algorithm, implemented in MEGA5, with search level 3 in which the initial trees were obtained with the random addition of sequences (10 replicates). The most parsimonious tree is shown (fig. 3). To estimate the statistical significance of the inferred clades in both analyses, 1,000 bootstrapping was performed to estimate the stability and support for the inferred clades.

Acknowledgements

We kindly acknowledge Jerry K. Davis (Department of Veterinary Pathobiology, Purdue University School of Veterinary Medicine, West Lafayette, Ind., USA) for providing culture stocks of *A. brassicae* and *A. palmae*.

COST-FA0807-STSM-180111-006678, the German Research Foundation (DFG) through projects KU 2679/2-1 and BU 890/21-1, the German Academic Exchange Service (DAAD) through project 56266384, the Ministry of Education and Science (Republic of Serbia, Grant No. TR31043) and the Max Planck Society supported this work.

References

Bai X, Correa VR, Toruno TY, Ammar el D, Kamoun S, Hogenhout SA: AY-WB phytoplasma secretes a protein that targets plant cell nuclei. Mol Plant Microbe Interact 2009;22:18–30.

Bai X, Zhang J, Ewing A, Miller SA, Jancso Radek A, Shevchenko DV, Tsukerman K, Walunas T, Lapidus A, Campbell JW, Hogenhout SA: Living with genome instability: the adaptation of phytoplasmas to diverse environments of their insect and plant hosts. J Bacteriol 2006;188:3682–3696.

Bischof DF, Vilei EM, Frey J: Functional and antigenic properties of GlpO from *Mycoplasma mycoides* subsp. *mycoides* SC: characterization of a flavin adenine dinucleotide-binding site deletion mutant. Vet Res 2009;40:35.

Brown DRB, Johansson JM, K-E: Genus I. *Acholeplasma* Edward and Freundt 1970; in Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB (eds): Bergey's Manual of Systematic Bacteriology. New York, Springer, 2011, vol 4, pp 688–696.

Carle P, Laigret F, Tully JG, Bove JM: Heterogeneity of genome sizes within the genus *Spiroplasma*. Int J Syst Bacteriol 1995;45:178–181.

Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Pujar A, Shearer AG, Travers M, Weerasinghe D, Zhang P, Karp PD: The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 2012;40:D742–D753.

Chernov VM, Chernova OA, Mouzykantov AA, Baranova NB, Gorshkov OV, Trushin MV, Nesterova TN, Ponomareva AA: Extracellular membrane vesicles and phytopathogenicity of *Acholeplasma laidlawii* PG8. ScientificWorldJournal 2012;2012:315474.

Connelly JC, Leach DR: The *sbcC* and *sbcD* genes of *Escherichia coli* encode a nuclease involved in palindrome inviability and genetic recombination. Genes Cells 1996;1:285–291.

Davis RE, Jomantiene R, Zhao Y: Lineage-specific decay of folate biosynthesis genes suggests ongoing host adaptation in phytoplasmas. DNA Cell Biol 2005;24:832–840.

de Crecy-Lagard V, El Yacoubi B, de la Garza RD, Noiriel A, Hanson AD: Comparative genomics of bacterial and plant folate synthesis and salvage: predictions and validations. BMC Genomics 2007;8:245.

Delcher AL, Harmon D, Kasif S, White O, Salzberg SL: Improved microbial gene identification with GLIMMER. Nucleic Acids Res 1999; 27:4636–4641.

Deutscher J, Francke C, Postma PW: How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. Microbiol Mol Biol Rev 2006; 70:939–1031.

Edengreen SJ, Markham PG: Multiplication and persistence of *Acholeplasma* spp in leafhoppers. J Invertebr Pathol 1987;49:235–241.

Fujita Y: Carbon catabolite control of the metabolic network in *Bacillus subtilis*. Biosci Biotechnol Biochem 2009;73:245–259.

Garriga X, Eliasson R, Torrents E, Jordan A, Barbe J, Gibert I, Reichard P: *nrdD* and *nrdG* genes are essential for strict anaerobic growth of *Escherichia coli*. Biochem Biophys Res Commun 1996;229:189–192.

Gon S, Faulkner MJ, Beckwith J: In vivo requirement for glutaredoxins and thioredoxins in the reduction of the ribonucleotide reductases of *Escherichia coli*. Antioxid Redox Signal 2006;8:735–742.

Gordon D, Abajian C, Green P: Consed: A graphical tool for sequence finishing. Genome Res 1998;8:195–202.

Hogenhout SA, Oshima K, Ammar el D, Kakizawa S, Kingdom HN, Namba S: Phytoplasmas: bacteria that manipulate plants and insects. Mol Plant Pathol 2008;9:403–423.

Hoischen C, Dijkstra A, Rottem S, Reizer J, Saier MH Jr: Presence of protein constituents of the Gram-positive bacterial phosphotransferase regulatory system in *Acholeplasma laidlawii*. J Bacteriol 1993;175:6599–6604.

Hoshi A, Oshima K, Kakizawa S, Ishii Y, Ozeki J, Hashimoto M, Komatsu K, Kagiwada S, Yamaji Y, Namba S: A unique virulence factor for proliferation and dwarfism in plants identified from a phytopathogenic bacterium. Proc Natl Acad Sci USA 2009;106:6416–6421.

Hudspeth DS, Vary PS: *spoVG* sequence of *Bacillus megaterium* and *Bacillus subtilis*. Biochim Biophys Acta 1992;1130:229–231.

Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, de Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, McMenamin C, Mi H, Mutowo-Muellenet P, Mulder N, Natale D, Orengo C, Pesseat S, Punta M, Quinn AF, Rivoire C, Sangrador-Vegas A, Selengut JD, Sigrist CJ, Scheremetjew M, Tate J, Thimmajanarthanan M, Thomas PD, Wu CH, Yeats C, Yong SY: Interpro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res 2012;40:D306–D312.

IRPCM: '*Candidatus* Phytoplasma', a taxon for the wall-less, non-helical prokaryotes that colonize plant phloem and insects. Int J Syst Evol Microbiol 2004;54:1243–1255.

Kall L, Krogh A, Sonnhammer EL: Advantages of combined transmembrane topology and signal peptide prediction – the Phobius web server. Nucleic Acids Res 2007;35:W429–W432.

Kube M, Beck A, Zinder SH, Kuhl H, Reinhardt R, Adrian L: Genome sequence of the chlorinated compound-respiring bacterium *Dehalococcoides* species strain CBDB1. Nat Biotechnol 2005;23:1269–1273.

Kube M, Migdoll AM, Gehring I, Heitmann K, Mayer Y, Kuhl H, Knaust F, Geider K, Reinhardt R: Genome comparison of the epiphytic bacteria *Erwinia billingiae* and *E. tasmaniensis* with the pear pathogen *E. pyrifoliae*. BMC Genomics 2010;11:393.

Kube M, Mitrovic J, Duduk B, Rabus R, Seemüller E: Current view on phytoplasma Genomes and encoded Metabolism. ScientificWorldJournal 2012;2012:185942.

Kube M, Schneider B, Kuhl H, Dandekar T, Heitmann K, Migdoll AM, Reinhardt R, Seemüller E: The linear chromosome of the plant-pathogenic mycoplasma '*Candidatus* Phytoplasma mali'. BMC Genomics 2008;9:306.

Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW: RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007;35:3100–3108.

Lazarev VN, Levitskii SA, Basovskii YI, Chukin MM, Akopian TA, Vereshchagin VV, Kostrjukova ES, Kovaleva GY, Kazanov MD, Malko DB, Vitreschak AG, Sernova NV, Gelfand MS, Demina IA, Serebryakova MV, Galyamina MA, Vtyurin NN, Rogov SI, Alexeev DG, Ladygina VG, Govorun VM: Complete genome and proteome of *Acholeplasma laidlawii*. J Bacteriol 2011;193:4943–4953.

Leach DR: Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. Bioessays 1994;16:893–900.

Lee IM, Davis RE, Gundersen-Rindal DE: Phytoplasma: phytopathogenic mollicutes. Annu Rev Microbiol 2000;54:221–255.

Lowe TM, Eddy SR: tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 1997; 25:955–964.

Maniloff J, Kampo GJ, Dascher CC: Sequence analysis of a unique temperature phage: mycoplasma virus L2. Gene 1994;141:1–8.

Marcone C, Neimark H, Ragozzino A, Lauer U, Seemüller E: Chromosome sizes of phytoplasmas composing major phylogenetic groups and subgroups. Phytopathology 1999;89: 805–810.

Marcone C, Seemüller E: A chromosome map of the European stone fruit yellows phytoplasma. Microbiology 2001;147:1213–1221.

Medrano-Soto A, Moreno-Hagelsieb G, Vinuesa P, Christen JA, Collado-Vides J: Successful lateral transfer requires codon usage compatibility between foreign genes and recipient genomes. Mol Biol Evol 2004;21:1884–1894.

Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KH, Remington KA, Anson EL, Bolanos RA, Chou HH, Jordan CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z, Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter JC: A whole-genome assembly of *Drosophila*. Science 2000;287: 2196–2204.

Nagai H, Kagan JC, Zhu X, Kahn RA, Roy CR: A bacterial guanine nucleotide exchange factor activates ARF on *Legionella* phagosomes. Science 2002;295:679–682.

Nagai H, Kubori T: Type IVb secretion systems of *Legionella* and other Gram-negative *Bacteria*. Front Microbiol 2011;2:136.

Neimark H, Kirkpatrick BC: Isolation and characterization of full-length chromosomes from non-culturable plant-pathogenic *Mycoplasma*-like organisms. Mol Microbiol 1993;7: 21–28.

Nishiyama K, Hanada M, Tokuda H: Disruption of the gene encoding p12 (*secG*) reveals the direct involvement and important function of SecG in the protein translocation of *Escherichia coli* at low temperature. Embo J 1994; 13:3272–3277.

Ogawa Y, Ooka T, Shi F, Ogura Y, Nakayama K, Hayashi T, Shimoji Y: The genome of *Erysipelothrix rhusiopathiae*, the causative agent of swine erysipelas, reveals new insights into the evolution of *Firmicutes* and the organism's intracellular adaptations. J Bacteriol 2011;193:2959–2971.

Oshima K, Kakizawa S, Nishigawa H, Jung HY, Wei W, Suzuki S, Arashida R, Nakata D, Miyata S, Ugaki M, Namba S: Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. Nat Genet 2004;36:27–29.

Rabus R, Kube M, Beck A, Widdel F, Reinhardt R: Genes involved in the anaerobic degradation of ethylbenzene in a denitrifying bacterium, strain EbN1. Arch Microbiol 2002;178:506–516.

Rabus R, Reizer J, Paulsen I, Saier MH Jr: Enzyme I(^{Ntr}) from *Escherichia coli*. A novel enzyme of the phosphoenolpyruvate-dependent phosphotransferase system exhibiting strict specificity for its phosphoryl acceptor, NPr. J Biol Chem 1999;274:26185–26191.

Razin S: The mycoplasmas. Microbiol Rev 1978;42:414–470.

Razin S, Yogev D, Naot Y: Molecular biology and pathogenicity of mycoplasmas. Microbiol Mol Biol Rev 1998;62:1094–1156.

Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B: Artemis: sequence visualization and annotation. Bioinformatics 2000;16:944–945.

Saccardo F, Martini M, Palmano S, Ermacora P, Scortichini M, Loi N, Firrao G: Genome drafts of four phytoplasma strains of the ribosomal group 16SrIII. Microbiology 2012;158:2805–2814.

Saito Y, Silvius JR, McElhaney N: Membrane lipid biosynthesis in *Acholeplasma laidlawii* b: de novo biosynthesis of saturated fatty acids by growing cells. J Bacteriol 1977;132:497–504.

Scherf U, Buckel W: Purification and properties of 4-hydroxybutyrate coenzyme A transferase from *Clostridium aminobutyricum*. Appl Environ Microbiol 1991;57:2699–2702.

Seemüller E, Sule S, Kube M, Jelkmann W, Schneider B: The AAA+ ATPases and HflB/FtsH proteases of '*Candidatus* Phytoplasma mali': phylogenetic diversity, membrane topology, and relationship to strain virulence. Mol Plant Microbe Interact 2013;26:367–376.

Shu HW, Liu TT, Chan HI, Liu YM, Wu KM, Shu HY, Tsai SF, Hsiao KJ, Hu WS, Ng WV: Genome sequence of the repetitive-sequence-rich *Mycoplasma fermentans* strain M64. J Bacteriol 2011;193:4302–4303.

Strauss E: Microbiology. Phytoplasma research begins to bloom. Science 2009;325:388–390.

Sugio A, MacLean AM, Kingdom HN, Grieve VM, Manimekalai R, Hogenhout SA: Diverse targets of phytoplasma effectors: from plant development to defense against insects. Annu Rev Phytopathol 2011;49:175–195.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011;28:2731–2739.

Tchieu JH, Norris V, Edwards JS, Saier MH Jr: The complete phosphotransferase system in *Escherichia coli*. J Mol Microbiol Biotechnol 2001;3:329–346.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997;25:4876–4882.

Toruno TY, Music MS, Simi S, Nicolaisen M, Hogenhout SA: Phytoplasma PMU1 exists as linear chromosomal and circular extrachromosomal elements and has enhanced expression in insect vectors compared with plant hosts. Mol Microbiol 2010;77:1406–1415.

Tran-Nguyen LT, Kube M, Schneider B, Reinhardt R, Gibb KS: Comparative genome analysis of '*Candidatus* Phytoplasma australiense' (Subgroup *tuf*-Australia I; *rp*-A) and '*Ca.* Phytoplasma asteris' strains OY-M and AY-WB. J Bacteriol 2008;190:3979–3991.

Tully JG, Whitcomb RF, Rose DL, Bove JM, Carle P, Somerson NL, Williamson DL, Eden-Green S: *Acholeplasma brassicae* sp. nov. and *Acholeplasma palmae* sp. nov., two non-sterol-requiring mollicutes from plant surfaces. Int J Syst Bacteriol 1994;44:680–684.

Vasconcelos AT, Ferreira HB, Bizarro CV, Bonatto SL, Carvalho MO, Pinto PM, Almeida DF, Almeida LG, Almeida R, Alves-Filho L, Assuncao EN, Azevedo VA, Bogo MR, Brigido MM, Brocchi M, Burity HA, Camargo AA, Camargo SS, Carepo MS, Carraro DM, de Mattos Cascardo JC, Castro LA, Cavalcanti G, Chemale G, Collevatti RG, Cunha CW, Dallagiovanna B, Dambros BP, Dellagostin OA, Falcao C, Fantinatti-Garboggini F, Felipe MS, Fiorentin L, Franco GR, Freitas NS, Frias D, Grangeiro TB, Grisard EC, Guimaraes CT, Hungria M, Jardim SN, Krieger MA, Laurino JP, Lima LF, Lopes MI, Loreto EL, Madeira HM, Manfio GP, Maranhao AQ, Martinkovics CT, Medeiros SR, Moreira MA, Neiva M, Ramalho-Neto CE, Nicolas MF, Oliveira SC, Paixao RF, Pedrosa FO, Pena SD, Pereira M, Pereira-Ferrari L, Piffer I, Pinto LS, Potrich DP, Salim AC, Santos FR, Schmitt R, Schneider MP, Schrank A, Schrank IS, Schuck AF, Seuanez HN, Silva DW, Silva R, Silva SC, Soares CM, Souza KR, Souza RC, Staats CC, Steffens MB, Teixeira SM, Urmenyi TP, Vainstein MH, Zuccherato LW, Simpson AJ, Zaha A: Swine and poultry pathogens: the complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae*. J Bacteriol 2005;187:5568–5577.

Vilei EM, Frey J: Genetic and biochemical characterization of glycerol uptake in *Mycoplasma mycoides* subsp. *mycoides* SC: its impact on H_2O_2 production and virulence. Clin Diagn Lab Immunol 2001;8:85–92.

Volokhov DV, Neverov AA, George J, Kong H, Liu SX, Anderson C, Davidson MK, Chizhikov V: Genetic analysis of housekeeping genes of members of the genus *Acholeplasma*: phylogeny and complementary molecular markers to the 16S rRNA gene. Mol Phylogenet Evol 2007;44:699–710.

Waack S, Keller O, Asper R, Brodag T, Damm C, Fricke WF, Surovcik K, Meinicke P, Merkl R: Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. BMC Bioinformatics 2006;7:142.

Wei W, Davis RE, Jomantiene R, Zhao Y: Ancient, recurrent phage attacks and recombination shaped dynamic sequence-variable mosaics at the root of phytoplasma genome evolution. Proc Natl Acad Sci USA 2008;105:11827–11832.

Worliczek HL, Kampfer P, Rosengarten R, Tindall BJ, Busse HJ: Polar lipid and fatty acid profiles – RE-vitalizing old approaches as a modern tool for the classification of mycoplasmas? Syst Appl Microbiol 2007;30:355–370.

Yeiser B, Pepper ED, Goodman MF, Finkel SE: SOS-induced DNA polymerases enhance long-term survival and evolutionary fitness. Proc Natl Acad Sci USA 2002;99:8737–8741.

Zhao Y, Wei W, Lee IM, Shao J, Suo X, Davis RE: Construction of an interactive online phytoplasma classification tool, iPhyClassifier, and its application in analysis of the peach X-disease phytoplasma group (16SrIII). Int J Syst Evol Microbiol 2009;59:2582–2593.